

## Comparison of Neuronal Networks with Extreme Gradient Boosting Methods for predicting departure-delays at Zurich Airport

Abflugverspätungen im Flugverkehr führen zu Ärgernissen der Passagiere und sind für Flughäfen sowie Airlines mit Kosten und Gewineinbussen verbunden. Genaue Prognosen können die Flughafen Zürich AG (Industriepartner dieser Arbeit) dabei unterstützen, frühzeitig und effizient auf zeitliche Änderungen in den Flughafenprozessen zu reagieren. Das Ziel dieser Arbeit ist es, neuronale Netze mit Extreme Gradient Boosting bei der Prognose von Verspätungsgründen und minutlichen Punktprognosen für Abflugverspätungen zu vergleichen.

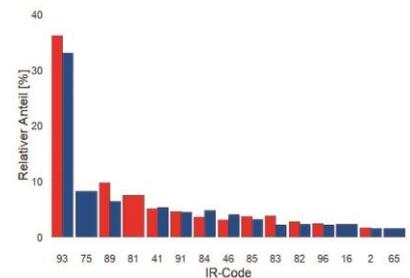
Für diese Arbeit liegen Daten der An- und Abflugpläne für den Zeitraum vom April 2015 bis August 2018 vor. Bei der Analyse werden zuerst die Abflugverspätungen auf verschiedene Zeitintervalle aggregiert und visualisiert, sodass ein genaueres Bild über die zeitlichen Entwicklungen entsteht. Anschliessend werden die Gründe, welche zu Verspätungen führen und in den Daten als Irregularity-Codes gespeichert sind, genauer analysiert. Es wird nach den Ursachen gesucht, welche am häufigsten auftreten und zu den meisten Verspätungen führen. Die Analyse zeigte, dass eine verspätete Ankunft des Vorgänger-Fluges (IR93) in den Winter- und Sommerflugplänen 2017 für ca. 20 % der verspäteten Abflüge sowie für ca. 35 % der gesamten Abflugverspätungen verantwortlich war. Auf-grund der Datengrundlage wurde entschieden, diese Art von Verspätung mittels einer binären Klassifikation zu prognostizieren. Die Variablen werden mittels Chi-Quadrat-Unabhängigkeitstests sowie einer logistischen Regression bestimmt und für die Punktprognose gleichermaßen verwendet.

Die Resultate von sieben getesteten Flügen zeigten, dass die Extreme Gradient Boosting-Methode in den meisten Fällen besser war als die neuronalen Netze. Extreme Gradient Boosting erreichte bei der Klassifikation eine durchschnittliche Genauigkeit von 95,1 % und bei der Punktprognose einen RMSE von 10.52 Minuten, während die neuronalen Netze eine Genauigkeit von 93,5 % und einen RMSE von 11.34 Minuten erreichten.

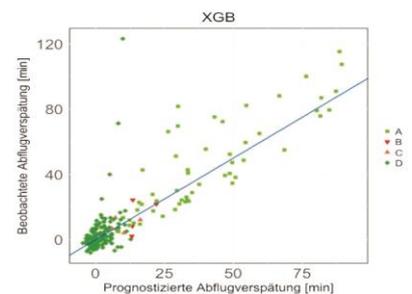


Diplomierende  
Cedric Käslin  
Reto Klauenbösch

Dozierende  
Manuel Renold  
Helmut Sedding



Der relative Anteil der IR-Codes an Verspätungen im Sommerflugplan (rot) und Winterflugplan (blau) 2017.



Resultate der Klassifikation und Punktprognose eines Fluges des Extreme Gradient Boosting. Die eingefärbten Kategorien stellen Klassifikationsergebnisse dar: A = TruePositive; B = FalsePositive; C = FalseNegative; D = TrueNegative